Published in final edited form as:
Hum Mutat. 2008 November ; 29(11): 1261–1264. doi:10.1002/humu.20903.

Assessing Pathogenicity: Overview of Results from the IARC Unclassified Genetic Variants Working Group

Sean V. Tavtigian[1,*], **Marc S. Greenblatt**[2], **David E. Goldgar**[3], and **Paolo Boffetta**[1] **for the IARC Unclassified Genetic Variants Working Group**†

[1]International Agency for Research on Cancer (IARC); Lyon, France

[2]Vermont Cancer Center and Department of Medicine, University of Vermont; Burlington, Vermont

[3]Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah

THE GROWING PROBLEM OF UNCLASSIFIED VARIANTS

Genetic testing for mutations in cancer susceptibility genes began in the early 1990 s and is now commonly performed [Frank et al., 2002; Lynch et al., 2007]. Important decisions regarding management of cancer susceptibility syndromes are made based on whether an individual carries a pathogenic variant or not. In addition to mutations that are highly likely to cause disease (deletions and truncations), mutation screening often finds missense substitutions, potential splicing variants, and/or small in-frame insertion-deletions that are of uncertain significance (often reported as unclassified variants [UVs] or [VUSs]). It was recognized early, and has come to pass, that unclassified missense substitutions would be a potentially serious problem unless functional assays or other discriminating tests could be established [Castilla et al., 1994]. For example, until recently, up to 12% of *BRCA1* and *BRCA2* test results reported a UV in the absence of any clearly pathogenic variant [Frank et al., 2002].

Over the last several years, research groups focused on individual diseases have begun to tackle this problem. However, to date, there has been little cross-fertilization. To foster cross-talk among gene-specific groups, and to organize and advance the field, the International Agency for Research on Cancer (IARC; the cancer research branch of the World Health Organization) convened a Working Group on Unclassified Sequence Variants in high-risk cancer susceptibility genes. The group met on February 4 and 5, 2008 and included investigators with specialties ranging from clinical cancer genetics to development of functional assays for susceptibility gene mutations (Table 1). The Working Group's goals were to establish standards for the approach to classification, including the terminology, evaluation, and validation of data types used in classification. This Special Issue of *Human Mutation* reports the discussions and recommendations of the Working Group in a series of seven articles. Three additional articles with related content are also presented.

DEGREE OF CERTAINTY OF PATHOGENICITY

The two most commonly tested genes are *BRCA1* and *BRCA2*. To date, the BRCA community, through the Breast Cancer Information Core (BIC) working group, has been conservative in classifying variants as pathogenic, consigning all variants with between 0.1% and 99%

[*]Correspondence to: Sean V. Tavtigian, International Agency for Research on Cancer; Lyon 69372, France. tavtigian@iarc.fr.
[†]The members of the Working Group are listed in the Appendix.

probability of pathogenicity as “unknown” [Goldgar et al., 2004]. For other genes, no such standards even exist, and conclusions of different laboratories and clinicians may differ. Quantifying the probability of pathogenicity of a variant can have important implications for clinicians who counsel patients regarding genetic testing results. A large number of variants do not reach the 99% level of confidence, and are ignored in making recommendations for management of cancer-susceptibility syndromes. This means that testing of at-risk relatives is not recommended, and that the individuals who carry such variants are treated the same as those who do not, even though there could be large differences in the cancer risk among them. Clinicians, however, are accustomed to making recommendations based on much higher degrees of uncertainty. It would be useful to define intermediate categories of risk to help make the difficult decisions inherent in cancer genetics, such as age and frequency of screening, or type and timing of surgical procedures. Many clinicians and patients might make different decisions based on a probability of 95% (or 5%) that a variant is pathogenic, rather than counseling on the basis of family history alone.

THE INTEGRATED METHOD

Clinicians regularly consider multiple sources of information to form a diagnostic conclusion. To classify genetic variants, there have been no widely accepted standards regarding which data should be considered and how to weigh their relative values. In October 2000, the BIC met to focus on the problem of variants in *BRCA1* and *BRCA2*. A result of the BIC meeting was a plan to develop a system that could combine data from several independent types of analysis to arrive at classifications of *BRCA1* and *BRCA2* variants. A series of works [Goldgar et al., 2004; Tavtigian et al., 2006; Chenevix-Trench et al., 2006; Easton et al., 2007; Spurdle et al., 2008b] analyzed many variants and eventually classified more than 150 with reasonable confidence.

The initial “integrated analysis” was based on likelihood ratios (LRs) that were generated independently for several statistical genetic methods. Each LR compared the probability of the observed data (e.g., cosegregation pattern) under the hypothesis that the variant was pathogenic (i.e., had the same effect as known truncating mutations) vs. the hypothesis that it was neutral with respect to cancer risk. These LRs were then multiplied together to achieve a final, integrated LR. The cutoffs that were assigned for pathogenicity and nonpathogenicity were conservative, with LR > 1,000 considered pathogenic and LR <0.01 considered neutral. All variants whose final LR fell between the cutoffs remained “uncertain.”

Since that time, the integrated evaluation has taken a more Bayesian approach, starting with some prior probability that a given variant is pathogenic and then modifying this probability by the observed data to arrive at a final (posterior) probability that the variant is deleterious. Details and additional discussion of this integration can be found in the accompanying article by Goldgar et al. [2008]. An empirically estimated prior probability is assigned, and is multiplied by LRs from independent analyses to arrive at a posterior probability of pathogenicity. This mechanism: 1) accommodates the uncertainty that accompanies all data types, as no analysis is likely to achieve predictive values of 100%; 2) automatically weights the contribution of individual methods—a method with an LR near 1.0 will contribute less toward a conclusion than a method with a large or small LR; and 3) provides a framework for adding new data types to the analysis as they are developed.

For most disease susceptibility genes, integrating data types to evaluate variants for pathogenicity, whether quantitatively or qualitatively, is a critical goal that has not yet been realized. To promote this area of study, the Working Group now presents a five-tiered classification system (pathogenic, likely pathogenic, uncertain, likely neutral, or neutral) for cancer susceptibility genes that links quantitative and qualitative data and clinical

recommendations. The system, as described in the accompanying article by Plon et al. [2008], accommodates uncertainty and the Working Group hopes that it can be extended to other disease susceptibility genes.

DATA SOURCES AND TERMINOLOGY

The data types available to be integrated can be divided into direct and indirect assessments of pathogenicity, and into those that provide quantitative vs. qualitative information. Because several measures available to researchers are actually making an indirect measurement of pathogenicity, we would like to clarify issues of terminology.

The three types of direct associations that have been quantified into LRs for or against pathogenicity of *BRCA1* or *BRCA2* variants are: 1) cosegregation with disease in pedigrees; 2) severity of personal and family cancer history; and 3) co-occurrence of a variant of interest with a clearly pathogenic variant in the same gene (summarized in Goldgar et al. [2008]). Tumor pathology methods can also be considered direct links of a variant and cancer, but they have not been quantified to determine the LR of pathogenicity of positive or negative results, as summarized by the accompanying article by Hofstra et al. [2008]. Indirect measurements include: 1) in silico assessments of sequence and structure variation based on evolutionary conservation [Tavtigian et al., 2008a, 2008b]; 2) assessment of a variant's potential effect on splicing [Spurdle et al., 2008a]; and 3) results from in vitro functional assays [Couch et al., 2008; Ollila et al., 2008]. Recent studies have begun to assign LRs to some in silico assessments, but LRs have not yet been applied to functional assays. Some useful functional assays (e.g., mismatch repair) could be used in qualitative classification schemes if standards could be adopted and there is a clear path toward validating their predictive values with sufficient accuracy to include their results in the quantitative "integrated evaluation." This cautious approach might extend to evaluation of mutations in genes of mendelian disorders that exhibit variable genotype/phenotype correlation; see Krasnov et al. [2008], as related to *CFTR*.

We encourage that the terms "pathogenic" and "nonpathogenic" be reserved only for variants for which multiple lines of evidence have been evaluated, or for which a convincing statistical association with disease is apparent based on large, sound studies of cases and controls. "Pathogenic" should not be applied to the conclusions of indirect forms of evidence, at least until such time as reliable estimates of the sensitivity, specificity, and predictive values of such assays are established through large multidisciplinary studies.

Sequence-based in silico measurements are really assessing a surrogate for pathogenicity, the effects of missense substitutions on evolutionary fitness. Following Kryukov et al. [2007], we propose to capture this distinction in our vocabulary by referring to variants that are likely to reduce evolutionary fitness as "deleterious." Likewise, functional assays also are assessing a surrogate, a missense substitution's effect on a specific protein function. Because many of the disease susceptibility genes that we are interested in are multifunctional, they may have measurable functions that are not important to susceptibility to the disease of interest. We propose to call variants that reduce in vitro protein function as "damaging" or "loss of function," and those that do not as "wild type" for that function.

WHO DOES THE CLASSIFICATION?

Major questions addressed in the accompanying database [Greenblatt et al., 2008; Ou et al., 2008] and clinical utility [Plon et al., 2008] articles are "who does the classifications?" and the related question of "how is the information disseminated to the relevant users?" In general:

- We DO expect panels composed of experts in all facets of a genetic syndrome (clinician, pathologist, molecular diagnostician, laboratory scientist, computational biologist, etc.) to do classification. Panels are likely but not guaranteed to include locus-specific database (LSDB) curators, and they may or may not include individuals from clinical testing labs. The process established by the BIC for BRCA variants and continued by the International Society for Inherited Gastrointestinal Tumors (InSiGHT) for mismatch repair variants should be used as models. The panels should be inclusive and transparent, and their conclusions should be clearly communicated to the relevant LSDBs and to testing laboratories. Perhaps a panel's updated conclusions could also be published in specialty journals that are likely to be read by the target audience concerned with the syndrome.
- We do NOT expect clinical testing labs to independently carry out a full integrated analysis and independently classify variants. We do expect them to gather data in a clinical report as best they can, and to faithfully report conclusions made by expert panels, in standardized terminology.
- We do NOT expect LSDB curators to carry out a full integrated analysis and independently classify variants. We do expect them to organize existing data and available classifications for public access through their LSDBs. The Summary Sheets provided on the BIC website are a prototype for other LSDB curators, in that they attempt to report all data and conclusions in a transparent manner.
- We do NOT expect physicians and or genetic counselors themselves to carry out a full integrated analysis and independently classify variants. We do expect them to: 1) know that data that they gather (e.g., medical and family histories) can help to modify existing classifications; 2) share the data with testing laboratories, LSDBs, and expert panels with appropriate patient consent; and 3) know how to present a fair summary of the data and conclusions to patients.

CONCLUSION

We hope that this Special Issue illustrates the progress, the barriers, and the future prospects in the field of classifying variants. By adhering to standards in terminology, data analysis, and reporting, researchers of variation in cancer susceptibility genes can establish clarity in this process and provide the benefits of improved interpretation to scientists, clinicians, and patients.

APPENDIX

MEMBERS OF THE IARC WORKING GROUP ON UNCLASSIFIED GENETIC VARIANTS

Paolo Boffetta, IARC, France; Fergus Couch, Mayo Clinic, USA; Niels de Wind, Leiden University, the Netherlands; Douglas Easton, Cambridge University, UK; Diana Eccles, University of Southampton, UK; William Foulkes, McGill University, Canada; Maurizio Genuardi, University of Florence, Italy; David Goldgar, University of Utah, USA; Marc Greenblatt, University of Vermont, USA; Robert Hofstra, University Medical Center Groningen, the Netherlands; Frans Hogervorst, Netherlands Cancer Institute, the Netherlands; Nicoline Hoogerbrugge, University Medical Center Neimejen, the Netherlands; Sharon Plon, Baylor University, USA; Paolo Radice, Istituto Nazionale Tumori, Italy; Lene Rasmussen, Roskilde University, Denmark; Olga Sinilnikova, Hospices Civils de Lyon, France; Amanda Spurdle, Queensland Institute of Medical Research, Australia; and Sean Tavtigian, IARC, France.

References

Castilla LH, Couch FJ, Erdos MR, Hoskins KF, Calzone K, Garber JE, Boyd J, Lubin MB, Deshano ML, Brody LC, Collins FS, Weber BL. Mutations in the BRCA1 gene in families with early-onset breast and ovarian cancer. Nat Genet 1994;8:387–391. [PubMed: 7894491]

Chenevix-Trench G, Healey S, Lakhani S, Waring P, Cummings M, Brinkworth R, Deffenbaugh AM, Burbidge LA, Pruss D, Judkins T, Scholl T, Bekessy A, Marsh A, Lovelock P, Wong M, Tesoriero A, Renard H, Southey M, Hopper JL, Yannoukakos K, Brown M, Easton D, Tavtigian SV, Goldgar D, Spurdle AB. Genetic and histopathologic evaluation of BRCA1 and BRCA2 DNA sequence variants of unknown clinical significance. Cancer Res 2006;66:2019–2027. [PubMed: 16489001]

Couch FJ, Rasmussen L, Hofstra R, Monteiro AANM, Greenblatt MS, de Wind N. IARC Unclassified Genetic Variants Working Group. Assessment of functional effects of unclassified genetic variants. Hum Mutat 2008;29:1314–1326. [PubMed: 18951449]

Easton DF, Deffenbaugh AM, Pruss D, Frye C, Wenstrup RJ, Allen-Brady K, Tavtigian SV, Monteiro AN, Iversen ES, Couch FJ, Goldgar DE. A systematic genetic assessment of 1,433 sequence variants of unknown clinical significance in the BRCA1 and BRCA2 breast cancer-predisposition genes. Am J Hum Genet 2007;81:873–883. [PubMed: 17924331]

Frank TS, Deffenbaugh AM, Reid JE, Hulick M, Ward BE, Lingenfelter B, Gumpper KL, Scholl T, Tavtigian SV, Pruss DR, Critchfield GC. Clinical characteristics of individuals with germline mutations in BRCA1 and BRCA2: analysis of 10,000 individuals. J Clin Oncol 2002;20:1480–1490. [PubMed: 11896095]

Goldgar DE, Easton DF, Deffenbaugh AM, Monteiro AN, Tavtigian SV, Couch FJ. Integrated evaluation of DNA sequence variants of unknown clinical significance: application to BRCA1 and BRCA2. Am J Hum Genet 2004;75:535–544. [PubMed: 15290653]

Goldgar DE, Easton DF, Byrnes GB, Spurdle AB, Iversen ES, Greenblatt MS. IARC Unclassified Genetic Variants Working Group. Integration of various data sources for classifying uncertain variants into a single model. Hum Mutat 2008;29:1265–1272. [PubMed: 18951437]

Greenblatt MS, Brody LC, Foulkes W, Genuardi M, Hofstra R, Plon S, Sijmons RH, Sinilnikova OM, Spurdle AB. Locus-specific databases and recommendations to strengthen their contribution to the classification of variants in cancer susceptibility genes. Hum Mutat 2008;29:1273–1281. [PubMed: 18951438]

Hofstra RW, Spurdle AB, Eccles D, Foulkes WD, de Wind N, Hoogerbrugge N, Hogervorst FBL. IARC Unclassified Genetic Variants Working Group. Tumor characteristics as an analytic tool for classifying genetic variants of uncertain clinical significance. Hum Mutat 2008;29:1292–1303. [PubMed: 18951447]

Krasnov K, Tzetis M, Cheng J, Guggino W, Cutting GR. Functional studies of rare missense mutations in CFTR facilitate interpretation of genotype-phenotype relationships. Hum Mutat 2008;29:1364–1372. [PubMed: 18951463]

Kryukov GV, Pennacchio LA, Sunyaev SR. Most rare missense alleles are deleterious in humans: implications for complex disease and association studies. Am J Hum Genet 2007;80:727–739. [PubMed: 17357078]

Lynch HT, Boland CR, Rodriguez-Bigas MA, Amos C, Lynch JF, Lynch PM. Who should be sent for genetic testing in hereditary colorectal cancer syndromes? J Clin Oncol 2007;25:3534–3542. [PubMed: 17687158]

Ollila S, Dermadi Bebek D, Jiricny J, Nystrom M. Mechanisms of pathogenicity in human MSH2 missense mutants. Hum Mutat 2008;29:1355–1363. [PubMed: 18951462]

Ou J, Niessen RC, Vonk J, Westers H, Hofstra RMW, Sijmons RH. A database to support the interpretation of human mismatch repair gene variants. Hum Mutat 2008;29:1337–1341. [PubMed: 18951442]

Plon SE, Eccles DM, Easton DF, Foulkes W, Genuardi M, Greenblatt MS, Hogervorst FBL, Hoogerbrugge N, Spurdle AB, Tavtigian S. IARC Unclassified Genetic Variants Working Group. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat 2008;29:1282–1291. [PubMed: 18951446]

Spurdle AB, Couch FJ, Hogervorst FBL, Radice P, Sinilnikova OM. IARC Unclassified Genetic Variants Working Group. Prediction and assessment of splicing alterations. Hum Mutat 2008a;29:1304–1313. [PubMed: 18951448]

Spurdle AB, Lakhani SR, Healey S, Parry S, Da Silva LM, Brinkworth R, Hopper JL, Brown MA, Babikyan D, Chenevix-Trench G, Tavtigian SV, Goldgar DE. Clinical classification of BRCA1 and BRCA2 DNA sequence variants: the value of cytokeratin profiles and evolutionary analysis—a report from the kConFab Investigators. J Clin Oncol 2008b;26:1657–1663. [PubMed: 18375895]

Tavtigian SV, Deffenbaugh AM, Yin L, Judkins T, Scholl T, Samollow PB, de Silva D, Zharkikh A, Thomas A. Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. J Med Genet 2006;43:295–305. [PubMed: 16014699]

Tavtigian S, Byrnes GB, Goldgar DE, Thomas A. Classification of rare missense substitutions, using risk surfaces, with genetic- and molecular-epidemiology applications. Hum Mutat 2008a;29:1342–1354. [PubMed: 18951461]

Tavtigian SV, Greenblatt MS, Lesueur F, Byrnes GB. IARC Unclassified Genetic Variants Working Group. In silico analysis of missense substitutions using sequence-alignment based methods. Hum Mutat 2008b;29:1327–1336. [PubMed: 18951440]

TABLE 1

WorkingGroup Participants

Participant	Primary field of interest	Board certification
Boffetta, Paolo	Molecular cancer epidemiology	
Byrnes, Graham B.	Statistical and genetics	NA
Couch, Fergus J.	Cancer genetics, molecular and cellular biology	NA
de Wind, Niels	Biochemistry of mismatch repair and in vitro MMR assays	NA
Easton, Douglas F.	Genetic epidemiology of cancer and statistical genetics	NA
Eccles, Diana M.	Clinical and molecular cancer genetics	Clinical genetics
Foulkes, William D.	Clinical and molecular cancer genetics	Medical genetics
Genuardi, Maurizio	Clinical and molecular cancer genetics	Hematology/oncology and medical genetics
Goldgar, David	Genetic epidemiology of cancer and breast cancer genetics	NA
Greenblatt, Marc S.	Medical oncology, cancer genetics, p16 assays, LSDBs	Internal medicine and medical oncology
Hofstra, Robert M.W.	Molecular genetics and gene functional characterization	NA
Hogervorst, Frans B.L.	Cancer molecular genetics, clinical mutation screening, LSDBs	lab: iso 17025
Hoogerbrugge, Nicoline	Clinical cancer genetics and clinical oncology	Internal medicine
Plon, Sharon E.	Clinical cancer genetics and molecular oncology	Clinical genetics
Radice, Paolo	Cancer genetics and molecular oncology	NA
Rasmussen, Lene J.	Basic molecular biology and biochemistry focused on MMR	NA
Sinilnikova, Olga	Cancer molecular genetics and clinical mutation screening	Medical genetics characterization
Spurdle, Amanda B.	Cancer molecular epidemiology	NA
Tavtigian, Sean V.	Cancer molecular genetics and bioinformatics	NA

[1]NA, not applicable.